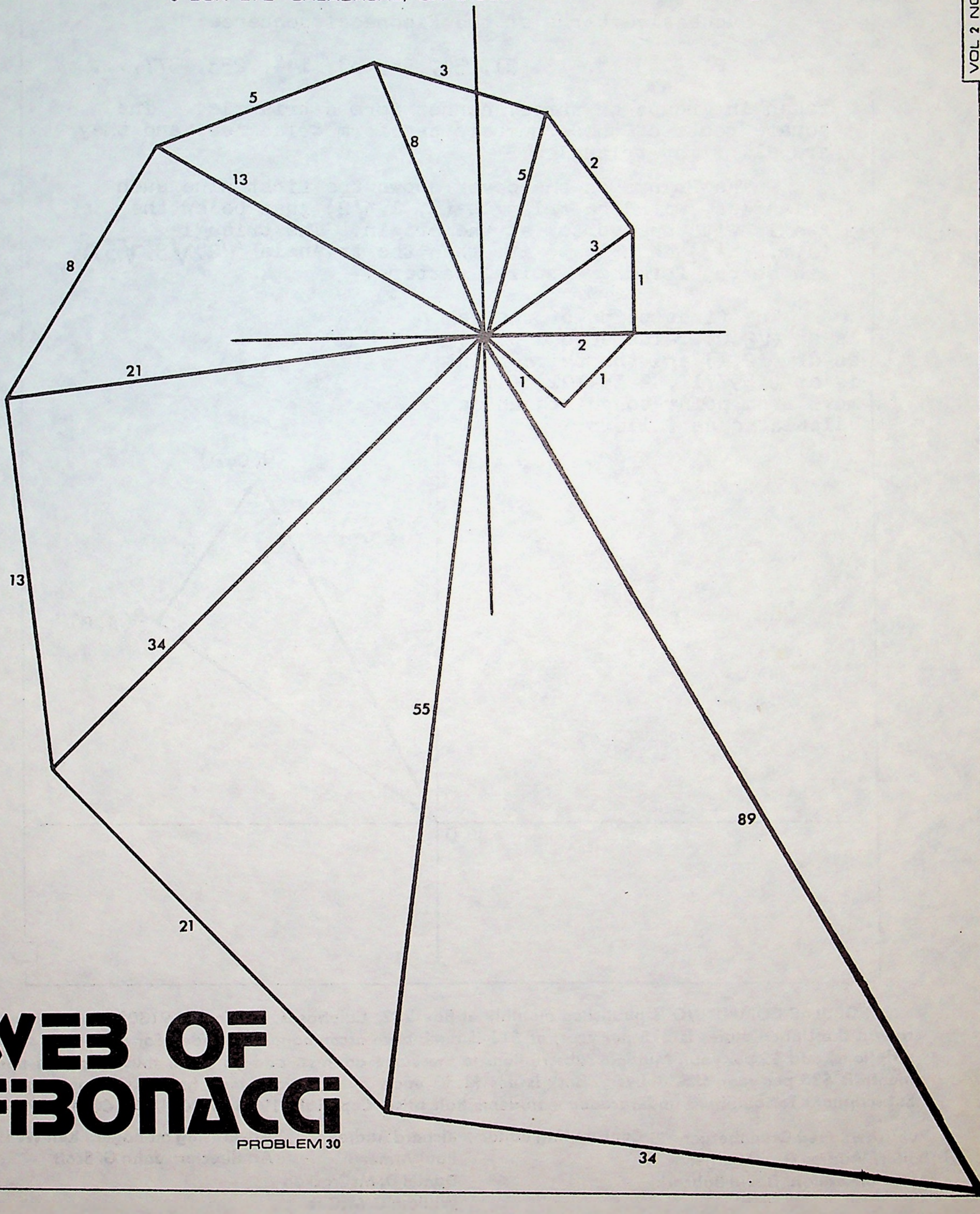
Popular Computing

BOX 272 CALABASAS, CA 91302

10

VOL 2 NO 1 JAN 74



WEB OF FIBONACCI

PROBLEM 30

THE WEB OF FIBONACCI

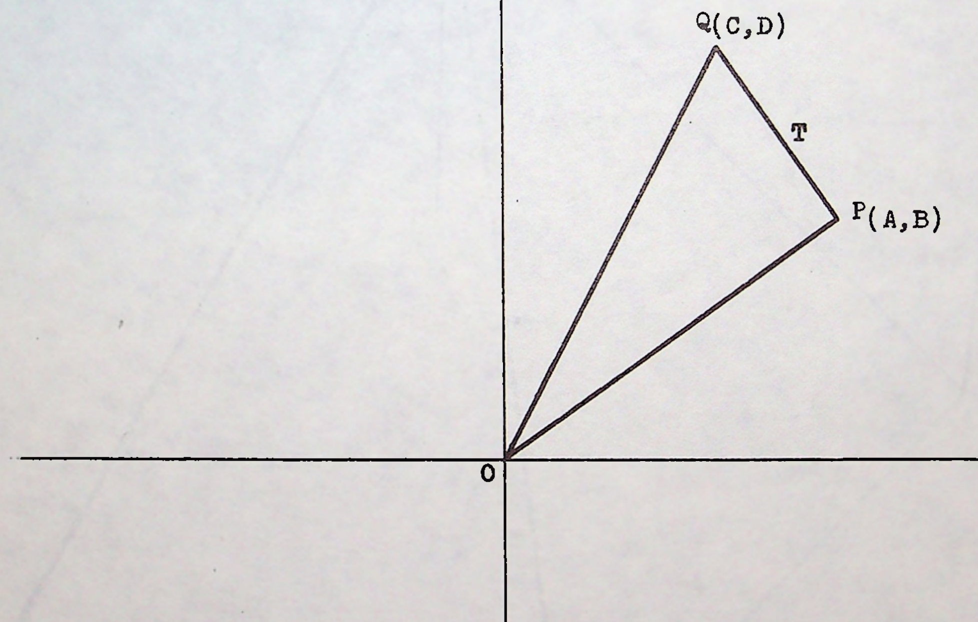
Successive terms of the Fibonacci sequence:

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377,...

taken in groups of three, cannot form a triangle. The square roots of these numbers can form triangles, and they are all right triangles.

The Figure on the cover shows the first nine such triangles, with the triangle $(1, 1, \sqrt{2})$ just below the x-axis with one vertex at the origin. The triangle $(1, \sqrt{2}, \sqrt{3})$ is next to it; then the triangle $(\sqrt{2}, \sqrt{3}, \sqrt{5})$, and so on, forming a spiral pattern.

The first point of the Web is at $(\sqrt{2}, 0)$. The second point is at $(\sqrt{2}, 1)$ and the third point is at (.597717, 2.154701). The move from point to point can be calculated as follows:



POPULAR COMPUTING is published monthly at Box 272, Calabasas, California 91302. Subscription rate in the United States is $15 per year, or $12 if remittance accompanies the order. For all foreign subscriptions, add $5 per year. Multiple subscriptions to the same address, add $5 each; thus, 3 copies per month is $25 per year, U.S. delivery. Back issues $1.50 each. Subscriptions may begin with any issue. Subscriptions for qualified undergraduate students half price. Copyright 1973 by POPULAR COMPUTING.

The slope of OP is B/A; the slope of PQ is thus -A/B. The equation of PQ is

$$\frac{y - B}{x - A} = -\frac{A}{B}$$

and Q must lie on this line. The square of the distance QP is

$$(D - B)^2 + (C - A)^2 = T^2$$

where T is the next term in the Fibonacci sequence. The two resulting equations in C and D can be solved by substitution, which leads to a quadratic equation. The root of this equation must be chosen so as to rotate the web counter-clockwise.

The Problem is: what are the coordinates of the 150th point of the Web?

POCKET CALCULATOR

Micro Instrumentation and Telemetry Systems (MITS) of Albuquerque offers a pocket calculator (model 941M) which, in addition to the arithmetic functions (8 digit floating point) has built-in constants for English-metric conversion, for lengths, volumes, areas, liquid measure, temperature, and mass. The machine sells in kit form for $130 or preassembled for $150.

BOUNDARIES

The numbers:

$$158733282881841916274491012923328901749236259319203520296443150620292 \pm 1$$

($76 \cdot 3^{139} \pm 1$) are the largest known twin primes, according to H. C. Williams and C. R. Zarmke, as given in Mathematics of Computation No. 120, October, 1972.

ART OF COMPUTING 1
·FLOWCHARTING·

Program flowcharts (as opposed to system flowcharts) form a topic that is older than computing itself. It is an emotion-charged topic, not only concerning how it should be done, but _why_. Respected experts in the field maintain that flowcharts are not necessary; that competent programmers can do their work without flowcharts. But even among those who agree that program flowcharts are a proper tool toward the construction of a tested program, there is little agreement on how flowcharts should be constructed. The only agreement is on the mechanics of flowchart construction, as defined in the standards promulgated by the International Standards Organization (ISO) and the American National Standards Institute (ANSI), X3.5 standard flowchart symbols. Consider some facts.

1. The first known computing flowchart was made by von Neumann, contained in a letter to Ed Paxon of the RAND Corporation in November 1946, some time before a true stored-program computer existed.

2. Flowcharting techniques were developed hit-or-miss by early workers in the field. In 1956, Russell McGee of General Electric wrote the first article on flowcharting (appearing in issue No. 76 of _Computing News_). McGee laid out a logical system, using only five symbols, and established that a program flowchart should concern itself only with the logic of the problem solution (that is, it should not involve details of any machine or any language. In other words, don't _code_ on a flowchart).

3. There is no shortage of material on the mechanics of flowcharting:

a. Ned Chapin, "Flowcharting With the ANSI Standard: A Tutorial," _Computing Surveys_, June 1970.

b. Ned Chapin, _Flowcharts_, Auerbach Publishing Company, 1971.

c. Mario Farina, _Flowcharting_, Prentice-Hall, 1970.

d. George Gleim, _Program Flowcharting_, Holt, Rinehart and Winston, 1970.

e. Thomas Schriber, _Fundamentals of Flowcharting_, Wiley, 1969.

f. IBM Corporation, _Flowcharting Techniques_, C20-8152, 1964.

g. M. Bohl, Flowcharting Techniques, Science Research Associates, 1971.

The ANSI standards should have settled some 20-year-old controversies, such as whether a decision box on a flowchart should be a diamond or an oval or whatever. As a matter of fact, textbooks continue to emerge using any symbols that the author favors. We do have a set of standards; we do not have any consistent adherence to it. But the standards still relate only to mechanics: what symbols to use; how to connect them; even the correct height-to-width ratio for a flowchart rectangle. Little is said (since McGee's 1956 article) on how to construct the logical part of a flowchart, or the purposes of flowcharts, or the things to avoid in flowcharting.

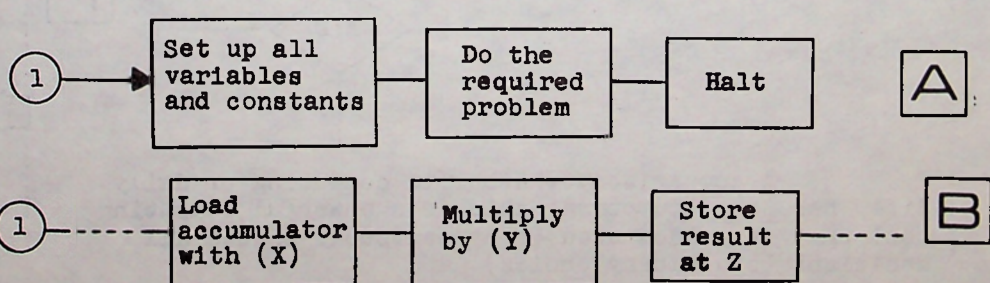
Consider first the purposes:

1. A flowchart is a graphic tool to help develop the logic of a problem solution.

2. A flowchart is a tool of communication. The communication may be from a programmer to others who will use (or patch) his program. Or, it may be from an author to his readers, as in a textbook. Most often, though, it is from a programmer to himself, at a later point in time, to clarify the logic that he has coded, in order to make corrections and changes to a working program.

It would be fatuous to deny that some people can write perfect programs, and maintain them, without using flowcharts. But experience shows that most beginners, in learning computing (that is, the coding phase) have great troubles with their logic if they do not use flowcharts. And, of course, the tale is repeated endlessly of the working program that needs modification by someone other than its author, which has to be completely reworked because the documentation (which should include the flowchart logic) is missing.

What should be the level of a flowchart? The highest possible level is shown at (A); the lowest level at (B). The proper level to solve the problem and help in making later modifications or corrections lies somewhere between these extremes. It's largely a matter of personal taste.

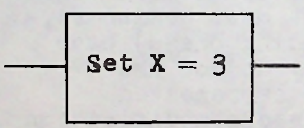


But then, almost everything about drawing flowcharts is a matter of personal taste. For example, there are "rules" that flowcharts should always be laid out vertically (or horizontally), but such rules are nonsense. Teachers tend to make them horizontally, since blackboards go that way; textbooks tend to the vertical layout, since book pages run vertically. Flowcharts produced by computer programs (e.g., Autoflow) run vertically, since paper flows vertically through a printer. There is no natural right way; if the purposes are met, then either way is correct. People seem to fall into the following non sequitur: "It really doesn't make any difference, but we ought to agree on a standard, and I learned it this way, so we should all agree on doing it my way." There is general agreement that a flowchart is to be read left-to-right and top-to-bottom like the printed page. When the direction of flow goes the other way, or is in any way ambiguous, arrowheads are used to keep things straight.

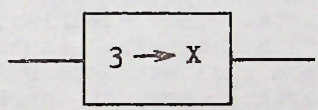
We can list some commonly accepted practices and some examples of poor practice.

1. Imperative commands are enclosed in rectangles; decisions are enclosed in diamonds.

2. All statements in flowchart boxes should be English sentences in some form. Thus, (C) specifies an imperative action, and (D) expresses precisely the same action, but uses the shorthand arrow to stand for "replaces."

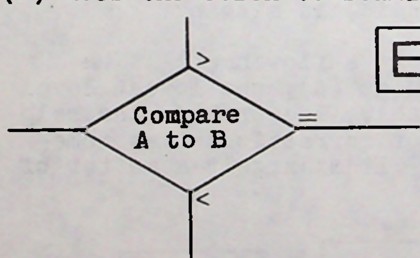


C

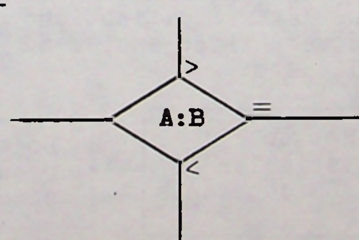


D

Similarly, (E) and (F) express the same decision, but (F) uses the colon to stand for "compare."



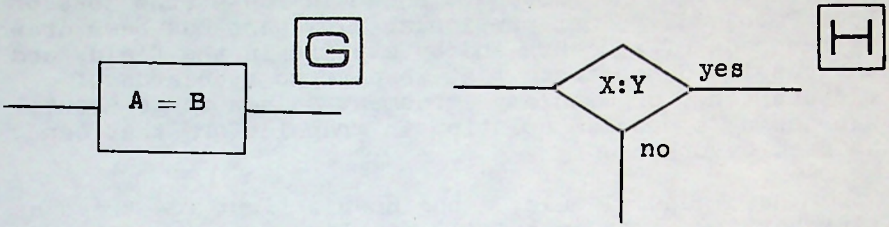
E



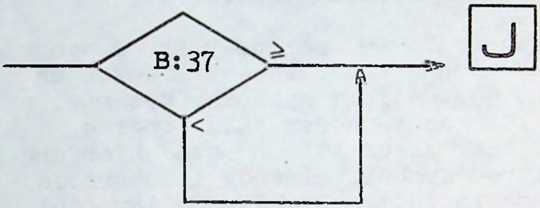
F

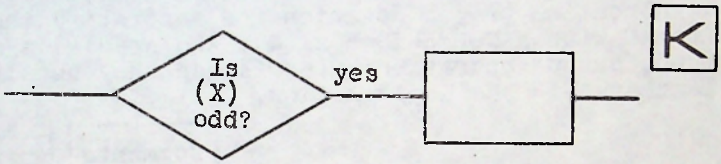
3. A comparison, whether in computing or daily life, has <u>three</u> outcomes; this is a powerful computing tool that is seldom used (many textbooks reduce all decisions to a binary choice).

Diagrams (G) and (H) illustrate illogical flow-charting. The statement in (G), although a correct English sentence, is ambiguous; it could mean "set A equal to B" or "set B equal to A" or "A and B are equal." The notation in (H) is simply not English.



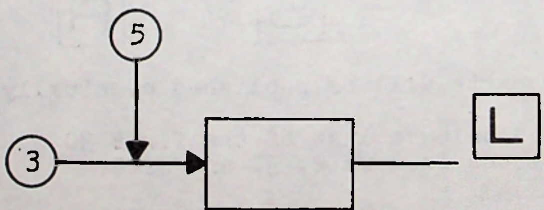
4. Two common (student) errors are illustrated in (J) and (K).

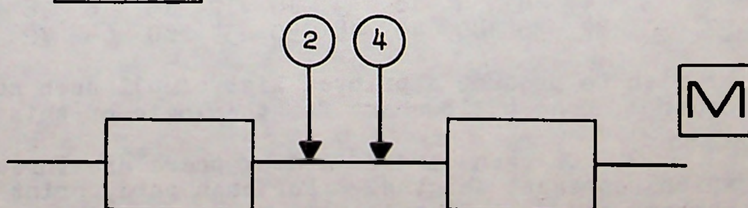




The illogic of (J)--both sides of a decision leading to the same point--is seldom so blatant as is shown, but shows up in clever disguises. (K) illustrates a checking procedure; namely, dangling ends on a flowchart are a sign of logical trouble.

5. Small circles are used as connectors, to avoid having to draw long connecting lines. There is nothing logically wrong with examples (L) and (M), but the usage is certainly redundant and frequently signifies muddled thinking.



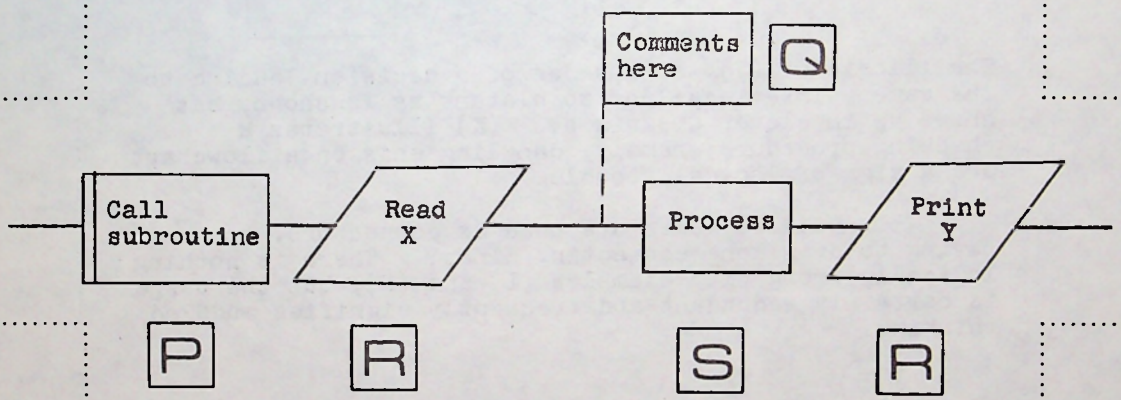


6. A flowchart will be easier to read if connecting lines do not cross. The intended direction of flow is hard to ascertain when the connecting lines cross.

The great difference between theory and practice is illustrated in the flowchart shown in PC9-9 (the Sets of Four Problem). That particular flowchart has been drawn by hundreds of students and by experts in the field, and the results demonstrate that we have no standards of flowcharting, or even any agreement on how to go about expressing a problem solution in graphic form that can be read by others.

As a general rule, a beginner's first few tries at flowcharting, even on trivial problems, are chaotic. It seems to be another one of those things that can be learned only by repeated and humbling experiences.

The flowchart shown in PC9-9 uses only three symbols: rectangles, diamonds, and connector circles. Texts on flowcharting offer up to 40 possible symbols, but most of them relate to system flowcharts rather than program flowcharts (and usually the reader is not told which are which). Besides the three symbols already illustrated, the most useful additions include a symbol for invoking a subroutine (P); a notation for annotating the flowchart itself with a COMMENTS box (Q); and special symbols for input/output operations, (R), since they usually constitute landmarks in the logical flow.



EXERCISES (sample flowcharts will be published eventually)

1. There is given below a list of the first 30 numbers containing only the factors 2, 3, and/or 5.

2	3	4	5	6	8	9	10	12	15	16	18	20	24	25
27	30	32	36	40	45	48	50	54	60	64	72	75	80	81

We wish to produce a printed list of all such numbers, in order. Draw a flowchart for the logic of this situation.

2. On each card of a deck there are three integers, which represent lengths. For each card, print the three numbers and an indication of the shape made by joining them. The possibilities are: no triangle, equilateral

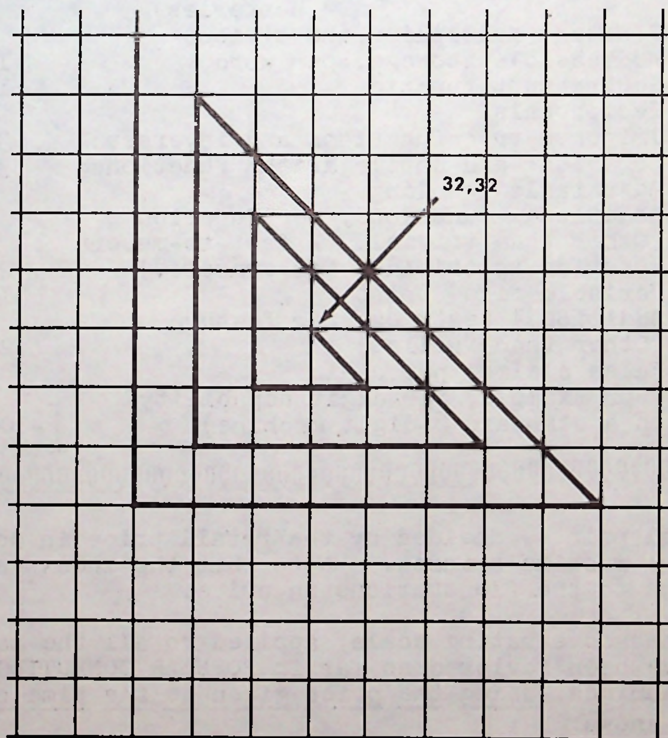
triangle, isosceles triangle, scalene triangle, and, for the last two cases, acute triangle or obtuse triangle.

3. Given in storage two integers, A and B. What is the logic of a subroutine whose calling sequence specifies the values A and B, and whose output should be the greatest common divisor of A and B?

4. The points on a 64 x 64 grid are to be scanned in the triangular "spiral" shown in (T), starting at point (32,32). That is, a subroutine is needed, for which the output will be the next point in the scan. The first few calls of the subroutine should produce the following:

32,32
33,31
32,31
31,31
31,32
31,33
31,34
32,33
33,32
34,31
35,30
34,30 and so on.

Draw a flowchart for this logic.



T

A Rating Scale for Calculators

With the profusion of brands and models of desk and pocket calculators now on the market, and with prices changing almost daily, a rating scale for such machines might be useful. Such a scheme is given below. It is not to be considered authoritative or even scientific. Anyone wishing to make single-number comparisons between two or more machines should apply his own weights to the features he regards as important.

In any event, we can assign point values to calculator features as follows:

1.	Floating point arithmetic	200
2.	Scientific notation	200
3.	Number of digits, D, in calculations	100(D-8)
4.	Number of digits, D, in display	20(D-8)
5.	AC operation only, no points	
	DC operation (i.e., portable, batteries)	50
	DC and AC (i.e., AC and rechargable batteries)	50
6.	Constant multiplier and divisor	50
7.	Addressable storage, per word	100
8.	Square root function	100
9.	Reciprocals	50
10.	Trigonometric functions and inverses	300
11.	Logarithm and antilogarithm functions	300
12.	Adjustable rounding	50
13.	Additional functions, per function (Other than trivial. A feet-to-meters function is trivial, for example.)	200
14.	Variable fixed point	100
15.	Additional features, per feature (Other than frills.)	200
16.	False claims, per claim (For example, "16-digit capability" on a standard 8-digit machine)	[-100]

The total points, divided by the retail price in dollars, gives the machine's index. Note that the index is thus sensitive to the fluctuations in price.

The above rating scale, applied to all the machines that have been reviewed so far in POPULAR COMPUTING, and other machines, using the price given at the time of the review, shows:

Model	Price	Index
Remington 6610	40	10.250
HP-80	395	6.683
TI SR-10	150	6.000
Compucorp 320	695	5.381
Craig 4505	120	5.000
HP-35	395	4.911
Victor 18-1721	495	3.575
Bowmar	100	3.500
Canon F-10	545	3.266
"Electronic Slideruler"	89	2.925
Dietzgen ESR-1	695	1.985

In almost every case, the index number for these machines is different now, since the prices have dropped. The TI SR-10, for example, now sells for under $100, which raises its index number to 9.000; the HP-35 at $295 has an index number of 6.576, and so on. Some of the basic machines (4 arithmetic functions, floating point, constant, battery operation, and that's all) can be obtained now for as low as $40, which would give them a rating of 8.750.

There are no points in the above ratings for subjective matters like

a. A well-known brand name.
b. A meaningful warranty.
c. Attractive layout of the keyboard.
d. An easily read display.
e. Long battery life and short recharging time.

Book Page Numbering PROBLEM 32

Book pages are consecutively numbered from 1 to K. Using individual pieces of type, how many of each of the decimal digits will be used to number the pages? Draw a flowchart for the logic involved. The input is K; the output is the distribution of the digits. For example, with input of 147, the output should be:

0	1	2	3	4	5	6	7	8	9
24	83	35	35	33	25	25	25	24	24

credibility game

PROBLEM 33

For any number of players. Each player moves a colored marker along the cells. Two honest dice are tossed for a move. A player's marker is advanced the number of cells indicated by the dice. If the move lands on one of the critical cells, it is moved back according to the stated directions.

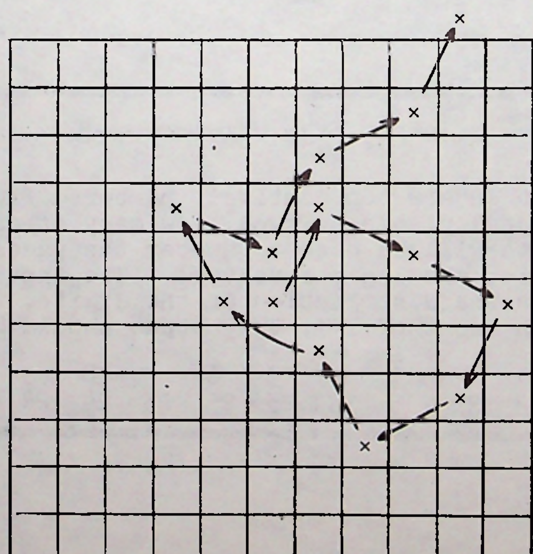
It is possible to reach or surpass cell 75 (thus winning the game) in seven moves (corresponding to dice tosses of 11, 12, 12, 12, 12, 12, and 4). A game may take many moves, but is unlikely to take more than 80 moves.

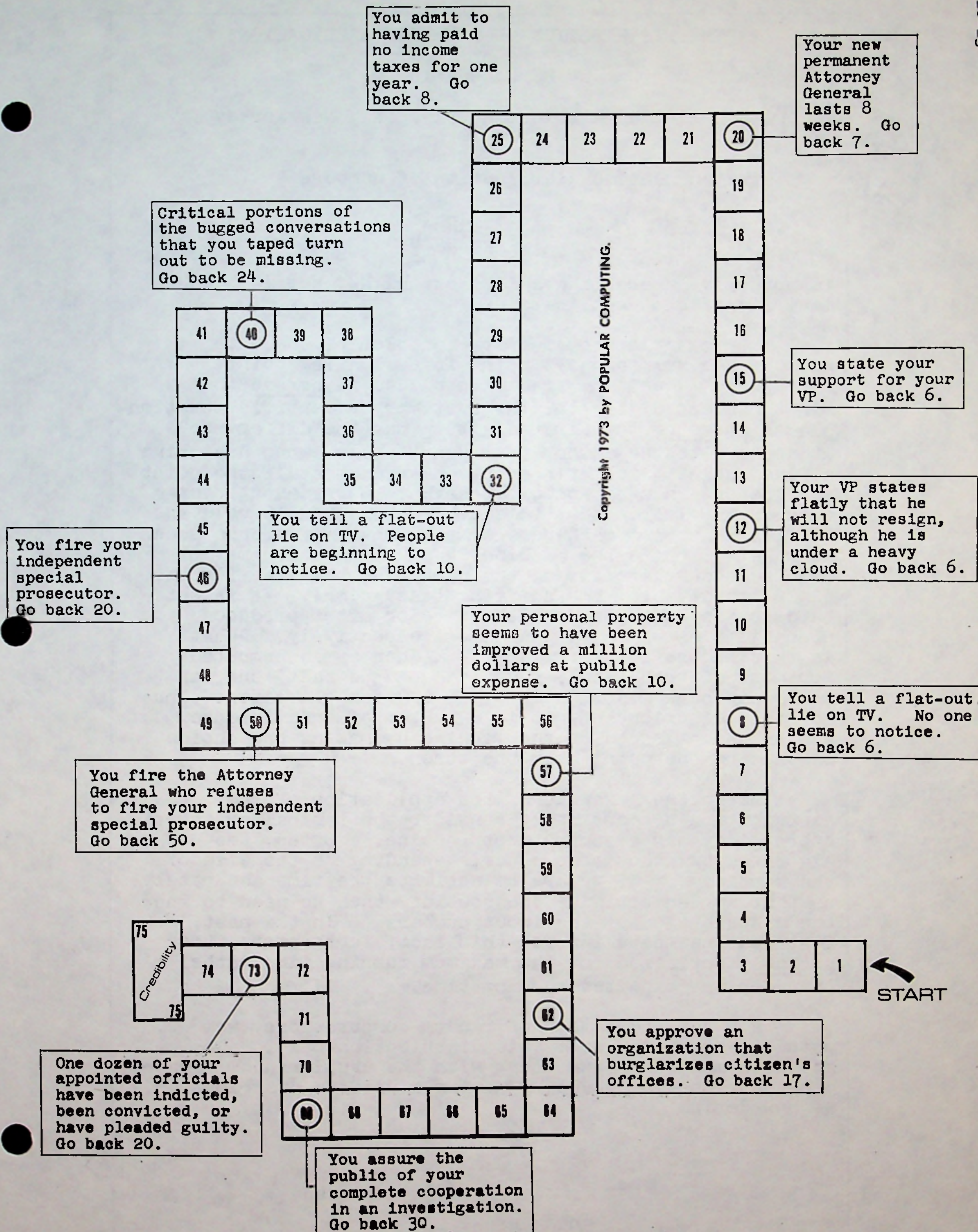
Problem: What is the distribution of the number of moves it takes to win the game?

A RANDOM WALK DISTRIBUTION

PROBLEM 34

A chess knight is at the center of an 11 x 11 board. The knight moves at random to one of the 8 squares possible, until the walk takes him off the board. What is the distribution of the lengths of the walks? A given trip can be as short as three legs, and would seldom exceed 20 legs. The Figure shows a walk of 12 legs.




START
Your new permanent Attorney General lasts 8 weeks. Go back 7.
You admit to having paid no income taxes for one year. Go back 8.
Critical portions of the bugged conversations that you taped turn out to be missing. Go back 24.
You state your support for your VP. Go back 6.
Your VP states flatly that he will not resign, although he is under a heavy cloud. Go back 6.
You tell a flat-out lie on TV. People are beginning to notice. Go back 10.
You fire your independent special prosecutor. Go back 20.
Your personal property seems to have been improved a million dollars at public expense. Go back 10.
You tell a flat-out lie on TV. No one seems to notice. Go back 6.
You fire the Attorney General who refuses to fire your independent special prosecutor. Go back 50.
Credibility
One dozen of your appointed officials have been indicted, been convicted, or have pleaded guilty. Go back 20.
You approve an organization that burglarizes citizen's offices. Go back 17.
You assure the public of your complete cooperation in an investigation. Go back 30.
Copyright 1973 by POPULAR COMPUTING.

THE DISTRIBUTION OF NUMBERS--APPLICATIONS

by R. W. Hamming

Bell Laboratories, Murray Hill, New Jersey

The cumulative distribution of numbers

$$D(x) = \frac{\log x}{\log b}$$

is not only interesting in its own right, but it has many significant applications.

For example, consider the problem of where to put the decimal (binary) point in a floating point computer. Computing tradition places it <u>before</u> the first nonzero digit, but the customary scientific notation places it <u>after</u> the first digit. Why the difference? It appears to the author that von Neumann, when designing a fixed point arithmetic unit, observed that if the point were before then he would not have to worry about overflow on the left. But this argument is hardly relevant to floating point machines. Well, what difference does it make? Inside the machine, a multiplication may require a shift to fix up the number so that the first digit is not zero, and shifting takes time. It is not hard to see that placing it before or after produces probabilities of shift p and 1-p respectively. What is this probability of a shift? When it is computed for most distributions, p depends on the base, but for the reciprocal distribution, $p = 1/2$ for all bases. Thus, for this distribution, it is a matter of indifference, and we cannot speed up the machine by making one choice of placing the point over the other.

As another example of the application of the distribution of mantissas, consider the typical arithmetic subroutine, say a square root routine. Often, the algorithm branches some place, depending on the size of the mantissa. If we are to estimate the time accurately so that we can optimize the routine, then we need to know the distribution of the input numbers. In the past, most routines have ignored this fact; furthermore, they <u>probably</u> optimized for the maximum running time rather than for the expected running time.

For the simulation of random computations, we should use numbers from this distribution D(x). To get such mantissas, we start with the usual random number generator that produces x_n having a uniform distribution. We now define

$$\left.\begin{aligned} Y_0 &= x_0 \\ Y_n &= x_n Y_{n-1} \end{aligned}\right\} \text{(shifted)}$$

(where "shifted" means shift, so that there are no leading nonzero digits). From the earlier results that numbers from a flat distribution when multiplied together rapidly approach $D(x)$ we conclude that the Y_n will have the distribution $D(x)$. Experimental verification on 8192 numbers showed this was remarkably true as judged by several statistical tests.

Roundoff estimation is a large field in which the distribution of the mantissas matters. Consider, for example, two numbers x_i $(i=1,2)$ with their associated errors e_i, and their product

$$\begin{array}{c} x_1 + e_1 \\ x_2 + e_2 \\ \hline x_1x_2 + x_1e_2 + x_2e_1 + e_1e_2 \end{array}$$

The main roundoff is usually from the terms

$$x_1e_2 + x_2e_1$$

and it is the leading digits of x_1 and x_2 that control the amount. But recall that small x_1 and x_2 increase the probability of a "shift left to remove the leading zero," and when this happens, you have

$$2(x_1e_2 + x_2e_1)$$

Thus, the "propagated error" is larger than expected on the average if the flat distribution is used. Generally speaking, the distribution $D(x)$ has been ignored in past studies of roundoff.

Again, if we are sorting <u>computed</u> <u>numbers</u>, then the distribution $D(x)$ may well be relevant to estimates of the "expected sorting time."

In the IBM 360/370 series, the number base is effectively 16. Thus, floating point numbers with mantissas of the form (in binary) can occur:

.1xxx...
.01xx...
.001x...
.0001...

It is easy to see that each type has probability 1/4 if the numbers are taken from the distribution $D(x)$, since

$$D(2^{-k}) - D(2^{-(k-1)}) = \frac{\log 2^{-k} - \log 2^{-(k-1)}}{\log 16} = 1/4$$

whereas the flat, uniform distribution gives very unequal probabilities. Thinking about this again reinforces our

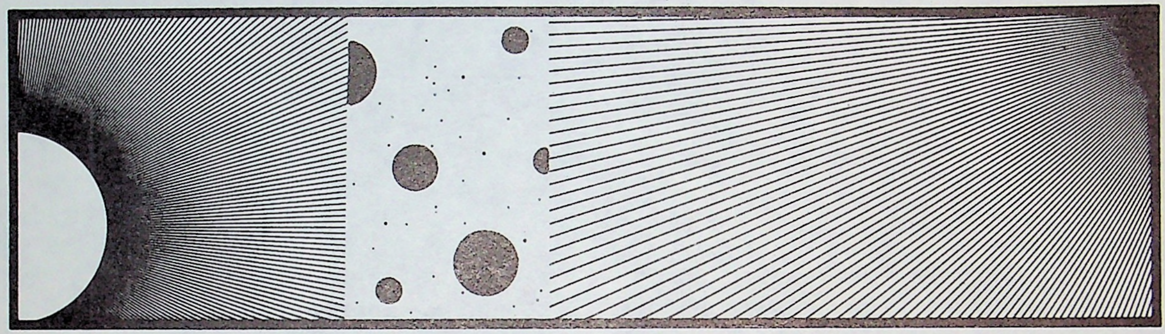
belief in the usefulness of the distribution D(x) in many situations.

EXERCISES

1. Simulate random products and examine the probability of a shift for (a) the flat, equilikely, distribution, and (b) the distribution D(x).

2. Construct and test a "random mantissa generator."

3. Discuss the construction of a least average running time square root routine.



N-SERIES

Ln 10	2.30258509299404568401799145468436420760110148862877 2976033327900967572609677352480235997205089598298 3
$\sqrt{10}$	3.16227766016837933199889354443271853371955513932521 6826857504852792594438639238221344248108379300295 2
$\sqrt[3]{10}$	2.15443469003188372175929356651935049525934494219211
$\sqrt[5]{10}$	1.58489319246111348520210137339150701326944213382504
$\sqrt[7]{10}$	1.38949549437313763712998521735301162211304671449100
$\sqrt[10]{10}$	1.25892541179416721042395410639580060609361740946693
$\sqrt[100]{10}$	1.02329299228075413096627517481987782734116405723798
e^{10}	22026.46579480671651695790064528424436635351261855 67 81074235426355225202818570792575199120968164 53
π^{10}	93648.04747608302097371669018491934563599815727551 47
$\tan^{-1} 10$	1.47112767430373459185287557176173085185530637718324
$\sqrt[10]{\pi}$	1.12128235323186329872203095522015520934269381296656
$\sqrt[10]{e}$	1.10517091807564762481170782649024666822454719473752